REMARKS/ARGUMENTS

1.) Claim Amendments

The Applicant has amended claims 1, 6-11, 15-23, 26, 29-33, and 37-44; and claim 28 has been canceled. Accordingly, claims 1-27 and 29-44 are pending in the application. Favorable reconsideration of the application is respectfully requested in view of the foregoing amendments and the following remarks.

2.) Claim Rejections – 35 U.S.C. § 103 (a)

The Examiner rejected claims 1-7, 17, 23-29 and 38 under 35 U.S.C. § 103(a) as being unpatentable over Cohen, et al. (US 5,825,771) in view of Taylor (US 5,982,305) and Kleijn (EP 0680033 A2). The Applicant has amended the claims. The Examiner's consideration of the amended claims is respectfully requested.

In contrast to Kleijn, the method according to the present invention is dynamic, i.e. could be altered from speech frame to speech frame. In previous Official Actions, the Examiner admits that EP-0680033 (Kleijn) teaches modifying the speech rate of the excitation signal provided by the speech signal source by either inserting or removing sample sequences of the excitation signal which correspond substantially to a pitch period (page 4, line 2-4). Thus, the solution according to Kleijn works on inserting or removing pitch periods, not on a sample-by-sample basis.

An argument put forward by the Examiner in the previous Official Action is that a pitch period could be compared to an individual sample. The Applicant respectfully disagrees. Normally, the pitch frequency is in the range from 50-60 Hz up to 400-500 Hz. e.g. most speech codecs use a pitch period in the interval 16-145 Hz, counted in the number of samples at a sampling frequency of 8 kHz. For example, in "Discrete-Time Processing of Speech Signals" (by Deller, Proakis and Hansen) it says in chapter 2.2.3 "Excitation of the Speech System and the Physiology of Voicing" on page 114 "For men, the possible pitch range is usually found somewhere between the two bounds 50-250 Hz, while for women the range usually falls somewhere in the interval 120-500 Hz." Consequently, if a pitch-frequency is 500 Hz, the pitch-period would be 16 samples (at 8 kHz sampling). If a pitch-frequency of is 50 Hz, the pitch-period would be 160 samples.

This characteristic is also used in speech coding with the purpose to not to waste bits on pitch periods that are impossible. For example, according to the speech coding standard GSM-EFR (ETSI 06.60) encoding and sending information is performed in pitch periods between 17.5 and 143. Other speech coding standards define other lowest and highest limits for possible pitch periods, but no existing speech coder has a lowest pitch period limit as low as 1 because this would be a waste of bits.

Because the claimed invention modifies the LPC-residual on a sample-by-sample basis, the play out buffer can be kept on a minimum and hence no extra delay is introduced. That is, an LPC-residual block comprising N samples is converted to a modified LPC-residual block comprising at least one of N-1 and N+1 samples, in response to a determination that a calculation, for each speech frame, so demands. The LPC-residual contains less information and less energy compared to the speech signal, but the pitch pulses P are still easy to locate. When modifying the LPC-residual, samples that are close to a pitch pulse P should be avoided, because these samples contain more information and thus have a large influence on the speech synthesis.

In Kleijn, where pitch periods are inserted or removed, it is assumed a fixed conversion factor between the transmitting and receiving side. Therefore, it cannot be used in dynamic systems, i.e. where the sampling frequency varies. Further, it does not solve the problem with under run or overrun situations, but is instead focused on changing the playback rate of a speech signal stored in compressed form for playback later and at a different speed to that at which it was stored. For example, in claim 1 of Kleijn it says: "predetermined change of SPEECH-RATE". According to the claimed invention, the original speech rate is the same, even if the sampling frequency would differ between the communicating terminals.

In US-5825771 (Cohen et al.), the number of packets/frames in the jitter buffer is the only parameter that has an influence on the decision whether packets/frames should be removed or added. The solution according to Cohen makes no difference due to which underlying problem causes the number of packets/frames in the jitter buffer to fluctuate.

There are mainly two problems that cause the number of packets/frames in the jitter buffer to fluctuate:

1) Delay jitter in the transmission

2) Different clock frequencies in the audio units

Buffering a sufficient number of packets/frames in the jitter buffer could solve the first problem. Cohen suggests an "all in one" solution to both problems.

The problem with handling both above identified problems in the same solution is that e.g. a difference in clock frequencies, that usually is fairly constant between two clients, is added to the delay jitter. The range (max Delay – min Delay) is often used as an input parameter to the adaptation algorithm that controls how many packets/frames to buffer in the jitter buffer. The added delay caused by e.g. the difference in clock frequencies in turn implies a delay for the users.

The users experience a "walkie-talkie"-like communication, which is very irritating for real time applications. Thus, the claimed invention aims to provide a solution for real time systems such as IP-telephony systems. In real time applications, it is necessary to keep the number of packets/frames in the jitter buffer on a low level, 1-3 packets/frames, which implies higher demands on jitter buffer adaptation.

Thus, the analysis in Cohen is done on a rough packet/frame by packet/frame level and therefore it takes a relatively long period of time before e.g. a difference in clock frequencies between two clients is detected. The adjustments (removing or adding packets/frames) in the jitter buffer according to Cohen are made on a rough packet/frame level, which will affect the speech quality negatively.

For real time applications such as IP-telephony, two separate solutions to the two above identified problems are required. The present invention is signified by fast and fine adjustments on a sample-by-sample basis, not on a rough packet/frame by packet/frame basis. Consequently, the claimed invention involves an inventive step over the prior art documents Kleijn and Cohen.

In US-5982305 (Taylor et al.), at least one zero sample is always inserted between each original sample in the interpolation phase. Taylor then extracts each and every N^{th} sample in the next decimation phase e.g. goes from 160 samples to 320 samples.

The steps according to Taylor:

1) Modify the sampling frequency (interpolation phase)

2) Apply a filter

3) Modify the sampling frequency (decimation phase)

Both the interpolation phase and the decimation phase works with integer numbers, which implies that only certain combinations are possible. This is the reason why the filter is applied to generate other combinations with sufficient quality, i.e. the decrease in quality is controlled.

The equivalent steps according to the claimed invention are:

1) LPC inverted filtering

2) Modify the length

3) LPC filtering

And the modification is only to add or remove one sample in a whole frame of samples, i.e. not in-between every sample.

So, Taylor does not add the elements of the claimed invention. Consequently, the claimed invention involves an inventive step over the prior art documents Kleijn, Cohen and Taylor.

The Examiner rejected claims 8 and 30 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Shlomot, et al. (EP 0680033). The Applicant has amended these claims. The Examiner's consideration of the amended claims is respectfully requested.

Shlomot has several differences from the claimed invention. For instance, when deciding if a correction should be made. Schlomot decides whether a correction should be made by adjusting the input jitter buffer (not the play out buffer) of the decoder, i.e. no decoding of the received speech frame has yet been performed when the adjustment is made.

In Schlomot, the decision is based on either the codec speech packet (CSP) (see claim 1 where the speech packet is compared with a threshold, or a parameter in the CSP, see claim 2 where a speech parameter of the CSP is compared with a threshold).

It should be noted that the CSPs are the packets that arrive at the decoder (the buffer is on the input to the decoder).

According to the claimed invention, the decision is triggered on other signals (see claim 5). It should be noted that the play out buffer listed in claim 5 contains the synthesized speech after the decoder (the buffer is on the output from the decoder).

In Schlomot, the frame/subframe/segment with low energy is found by comparing a parameter in the CSP with a threshold, see claim 2. A slightly more detailed description is found on page 5, column 8, rows 17-22 where it is described that the energy and spectral slope parameters are used and that these are transmitted together with the other information in the CSPs.

In the amended independent claims, the segment is found in the following way:

1) decode the received parameters and create a (first and temporary) excitation frame

2) analyze the (first and temporary) excitation frame to find a segment with low energy

3) modify the (first and temporary) excitation frame to create a second excitation

4) use the second (and modified) excitation frame when filtering with the LPC filter to create the synthesized speech

In Schlomot, it is actually the encoder that finds the segment since it is the encoder that calculates the energy, and possibly also the spectrum tilt, of a frame/subframe/segment. The encoder also encodes the energy (and spectrum tilt) information and transmits at least one parameter. The only thing that the decoder does is to take the energy and/or spectrum tilt parameter(s) and compare it with one (or several) threshold(s). This means that the method that Schlomot proposes only works if the energy and/or spectrum tilt parameter(s) are explicitly transmitted together with the other parameters in the coded speech frame.

In the claimed invention, in contrast, the low energy segment is found by analyzing the reconstructed excitation generated in the decoder. This method is preferred since most modern speech codecs do not send any explicit energy or spectrum tilt information. They usually send gain parameters for the adaptive (LTP) and

fixed code book excitations and resonance frequencies for the LPC filter, but one has to reconstruct the excitation before one can analyze it and find the low energy segment.

In Schlomot, one may wait until the clock skew has accumulated until the modification is made. This is shown on page 5, column 8, row 25-26 where it is described that one can wait some time before applying the corrections, and also in row 37-45 where one first allows corrections only in silence frames and then (after some frames) modifies the threshold so that corrections can also be made during unvoiced frames, and if this is not possible one can force a correction anyway.

In the claimed invention, a correction is made immediately, within the same frame, if it is decided that a correction should be made.

The problem with Schlomot's method is that the accumulated timing error, due to clock skew, cannot be distinguished from transmission delay jitter. It may not look like a big deal, but if the jitter buffer level, for the buffer that holds the CSPs, is made adaptive, then this timing error will introduce an error in the statistical analysis of the delay jitter, and then it is very likely that the adaptation will choose to apply a jitter buffer level that is unnecessarily high.

In the claimed invention, the clock skew is separated from any delay jitter, since the error, due to clock skew, is compensated for continuously. Thereby, the uncertainty in the statistical analysis for the jitter buffer level adaptation is smaller with the claimed invention.

Thus, even if elements from Schlomot were combined with Cohen, the combination would not produce a workable invention. Thus, a combination of Schlomot and Cohen does not render the claimed invention obvious.

The Examiner rejected claims 9 and 31 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Galand, et al. (US 5,073,938). The Applicant has amended the claims to better distinguish the claimed invention from Cohen. Claims 9 and 31 depend from amended claim 1 and 23 and recite further limitations in combination with the novel elements of claims 1 and 23. Galand does not provide the limitations of the amended base claims. Therefore, the allowance of claims 9 and 31 is respectfully requested for the same reasons that claims 1 and 23 are allowable.

The Examiner rejected claims 10 and 32 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Graumann, et al. (US 5,598,466). The Applicant has amended the claims to better distinguish the claimed invention from Cohen. Claims 10 and 32 depend from amended claim 1 and 23 and recite further limitations in combination with the novel elements of claims 1 and 23. Graumann does not provide the limitations of the amended base claims. Therefore, the allowance of claims 10 and 32 is respectfully requested for the same reasons that claims 1 and 23 are allowable.

The Examiner rejected claims 11, 12, 33 and 34 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Arjmand, et al. (US 5,067,158). The Applicant has amended the claims to better distinguish the claimed invention from Cohen as modified by Taylor and Kleijn. Claims 11, 12, 33 and 34 depend from amended claim 1 and 23 and recite further limitations in combination with the novel elements of claims 1 and 23. Arjmand does not provide the limitations of the amended base claims. Therefore, the allowance of claims 11, 12, 33 and 34 is respectfully requested for the same reasons that claims 1 and 23 are allowable.

The Examiner rejected claims 13 and 35 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Galand. The Applicant has amended the claims to better distinguish the claimed invention from Cohen as modified by Taylor, Kleijn, and Arjmand. Claims 13 and 35 depend from amended claim 1 and 23 and recite further limitations in combination with the novel elements of claims 1 and 23. Galand does not provide the limitations of the amended base claims. Therefore, the allowance of claims 13 and 35 is respectfully requested for the same reasons that claims 1 and 23 are allowable.

The Examiner rejected claims 14 and 36 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Galand. The Applicant has amended the claims to better distinguish the claimed invention from Cohen as modified by Taylor

and Kleijn. Claims 14 and 36 depend from amended claim 1 and 23 and recite further limitations in combination with the novel elements of claims 1 and 23. Galand does not provide the limitations of the amended base claims. Therefore, the allowance of claims 14 and 36 is respectfully requested for the same reasons that claims 1 and 23 are allowable.

The Examiner rejected claims 15, 16, 17, 37, 38 and 39 under 35 U.S.C. § 103(a) as being unpatentable over Cohen and further in view of Lee, et al. (US 5,617,507). The Applicant has amended the claims to better distinguish the claimed invention from Cohen as modified by Taylor and Kleijn. Claims 15, 16, 17, 37, 38 and 39 depend from amended claim 1 and 23 and recite further limitations in combination with the novel elements of claims 1 and 23. Lee does not provide the limitations of the amended base claims. Therefore, the allowance of claims 15, 16, 17, 37, 38 and 39 is respectfully requested for the same reasons that claims 1 and 23 are allowable.

CONCLUSION

In view of the foregoing remarks, the Applicant believes all of the claims currently pending in the Application to be in a condition for allowance. The Applicant, therefore, respectfully requests that the Examiner withdraw all rejections and issue a Notice of Allowance for all pending claims.

The Applicant requests a telephonic interview if the Examiner has any questions or requires any additional information that would further or expedite the prosecution of the Application.

Respectfully submitted,

Date: 6-14-04

Bill R. Naifeh
Registration No. 44,962

Ericsson Inc.
6300 Legacy Drive, M/S EVR1 C-11
Plano, Texas 75024

(972) 583-2012
bill.xb.naifeh@ericsson.com